References

[1] *Structure of microtubules* (Eds K.Roberts and J.S Hyams) (Academic Press).

[2] Gittes, F., Mickey, B., Nettleton, J. and Howard, J., *J. of Cell Biol.* (1993) **120(4)**, 923-934.

[3] Venier, P., Maggs, A.C., Carlier, M.F. and Pantaloni, D., *J. Biol. Chem.* (1994) **269(18)**, 13353-13360.

[4] Pirollet, F., Job, D., Margolis, R. and Garel, J., *EMBO journal* (1987) **6(11)**, 3247-3252.

[5] Mandelkow, E., *Methods in Enzymology* (1986) **134**, 149-168.

[6] Suzuki, A., Maeda, T. and Ito, T., *Biophys. J.* (1991) **59**, 25-30.

[7] Flory, P.J., *Proc. Royal. Soc. London* (1956) **A234**, 73-89.

[8] Bras, W., Diakun, G.P., Díaz, J.F., Maret, G., Kramer, H., Bordas, J. and Medrano, F.J., *Biophysical Journal* (1998) **74(3)**, 1509-1521.

[9] http://www.dl.ac.uk/SRS/CCP13

[10] Denny, R.C., CCLRC Daresbury laboratory private communication.

[11]Klug, A., Crick, F. and Wyckoff, H., *Acta. Cryst.* (1958) **11**, 199-213.

[12] Amos, L.A. and Klug, A., *J. Cell. Sci.* (1974) **14**, 523-549.

[13] Andreu, J.M., Bordas, J., Díaz, J.F., Garcia de Ancos, J., Gil, R., Medrano, F.J., Nogales, E., Pantos, E. and Towns-Andrews, E., *J. Mol. Biol.* (1992) **226**, 169-184.

[14] Glatter, O., in *Small Angle X-ray scattering* (Eds Glatter, O. and Kratky, O.) (Academic Press, 1982).

[15] Vainshtein, B.K., in *Diffraction of X-rays by chain molecules* (Elsevier, 1966).

[16] Edwards, S.F., *Proc. Phys. Soc.* (1966) **88**, 265-281.

* A part of the problem due to the poor orientation might be overcome by keeping the samples during the experiments in a magnetic field. Preparations, in collaboration with the High Magnetic Field Laboratory (Grenoble), to use a 10 T split coil magnet are in an advanced state. Interested people can contact W. Bras.

The ups and downs of native cellulose structure

R.P. Millane

Whistler Center for Carbohydrate Research and Computational Science and Engineering Program, Purdue University, West Lafayette, Indiana 47907-1160, U.S.A.

Cellulose was one of the first materials to be studied by X-ray fibre diffraction. Although structural studies have continued over the last 60 years, ambiguities have persisted with regard to chain packing and multiple phases of native cellulose, and only recently have these been resolved. X-ray fibre diffraction has played a dominant role, although other techniques have also provided essential information.

Cellulose is the major constituent of most land plants, is the most abundant natural compound, and is an important commercial raw material. It is a linear polymer of 1→4-linked β-D-glucose with a degree of polymerisation >3000. In native cellulose, the molecules are aligned to form fibres, some regions of which have an ordered crystalline structure. The crystalline regions vary in size, are mechanically strong, and are resistant to chemical and enzymatic attack. Cellulose is a structural component in plant and other systems and is used widely in industry. Artificial cellulose derivatives are also used extensively in a variety of industries. Although plant sources are the most familiar, cellulose is also present in bacteria, fungi and algae. Due to its ubiquity and importance, and its polycrystalline nature, cellulose was one of the first materials to be studied by X-ray fibre diffraction analysis. The first such studies were reported by Meyer and Misch [1] in 1937, on plant cellulose from ramie. They determined that the unit cell is monoclinic, that two molecules pass through the unit cell, and suggested that the chains have two-fold screw symmetry. Further studies, including data from electron diffraction and infrared spectroscopy, and also from algal celluloses, supported the results of Meyer and Misch, but with some differences [2-4]. Most notably, diffraction patterns from the algal celluloses contained a few, extra, weak reflections. These were attributed to the *a* and *b* unit cell dimensions being twice those of the plant celluloses. This "large" unit cell therefore contains eight molecules, and there were presumably small differences between the packing and/or molecular

conformations of the chains that produce small deviations from the symmetry of the small two-chain unit cell. Interestingly (this will be discussed later), Sarko and Muggli [3], in 1974, noted that the equatorial reflections (that include some of the "extra" large unit cell reflections) on an electron diffraction pattern from a bacterial cellulose could all be indexed on the basis of a one-chain triclinic unit cell.

The above discussion refers to "native" cellulose, which occurs only as a result of biosynthesis, and is referred to as cellulose I. Although this article is concerned only with native cellulose, we note, for completeness, that various treatments of cellulose produce a variety of polymorphic crystal structures [5]. These all retain the same molecular structure of native cellulose, but differ in the crystal packing and the intermolecular hydrogen bonding pattern. The different allomorphs are referred to as cellulose II, III, and IV. Cellulose II is produced by mercerization (swelling in alkali) or regeneration (precipitation) of cellulose I, and has a two-chain monoclinic unit cell, but of different cell dimensions to those of cellulose I. The cellulose I $\rightarrow$ II transition can be conducted in the solid state, is irreversible, and corresponds to a parallel to antiparallel transition in the packing. There are two forms of cellulose III that are obtained by treating cellulose I or II with liquid ammonia, and the molecules pack in two different two-chain monoclinic unit cells of (again) different dimensions. Two forms of cellulose IV are obtained by heating cellulose I or II with gylcerol, and the molecules pack in two different two-chain orthorhombic unit cells.

Returning to cellulose I, two chains pass through the monoclinic unit cell at the points (0,0) and (1/2,1/2) in the *a*-*b*-plane, and the space group is P1. Within these constraints, there are three possible packing arrangements of the two chains. First, the chains may be "parallel" or "antiparallel," as a result of the directionality of the cellulose molecule. Second, because of the monoclinic angle, there are two different parallel packings, that are referred to as "parallel up" and "parallel down". That there are two distinct parallel packings appears to have been first recognised by Gardner and Blackwell [4]. Because the monoclinic angle $\gamma = 97°$ is close to 90°, the differences between the up and down packings are small, but are stereochemically significant.

The first quantitative (i.e. based on an objective numerical refinement of crystal structure models against X-ray and steric data) X-ray structures of cellulose I were by Sarko and Muggli [3] and Gardner and Blackwell [4], both in 1974, for *Valonia* cellulose. These are referred to here as the SM and GB *Valonia* structures, respectively. In both of these studies, the weak "large cell" reflections were excluded from the analyses and the structure determined on the basis of the small one-chain unit cell. Following these studies, in 1980 Woodcock and Sarko [6] determined the structure of native ramie cellulose (referred to here as the WS ramie structure). These three structures emerged as the definitive crystal structures of native cellulose. They consistently defined a ribbon-like 2_1 structure for the cellulose molecule, two of which packed parallel in the unit cell, and the formation of hydrogen-bonded sheets. However, they were inconsistent regarding the packing (up or down) of the sheets, which are stabilised by van der Waals forces. Unfortunately, this was not generally realised at the time since the differences (1) are rather small, and (2) were confused by the use of differing conventions for describing the packings in the different studies. The confusion was due to the SM and GB *Valonia* structures both being described as parallel up, whereas, in fact, they actually correspond to opposite packing polarities. Although the same definitions of "up" packing of the chains relative to the *c*-axis was used, the unit cell constants were defined such that $a < b$ for the SM *Valonia* structure, whereas the GB structure was defined with $a > b$. This effected a reversal in the polarities between the two structures. In what has now become the standard convention ($a < b$), the SM structure is parallel up, whereas the GB structure is parallel down. This inconsistency went largely unnoticed until it was pointed out by French in 1989 [7]. Furthermore, Sarko and Muggli did not, in fact, consider the parallel down packing in their analysis. The WS ramie cellulose I structure is packed parallel up. There was also, therefore, a question as to whether the plant and algal celluloses have the same, or different, packings. Consequently, even in the 1990s, no definitive X-ray structure of *Valonia* cellulose I actually existed. Furthermore, since *Valonia* gives the best X-ray data, the precise packings of all the native celluloses were in doubt.

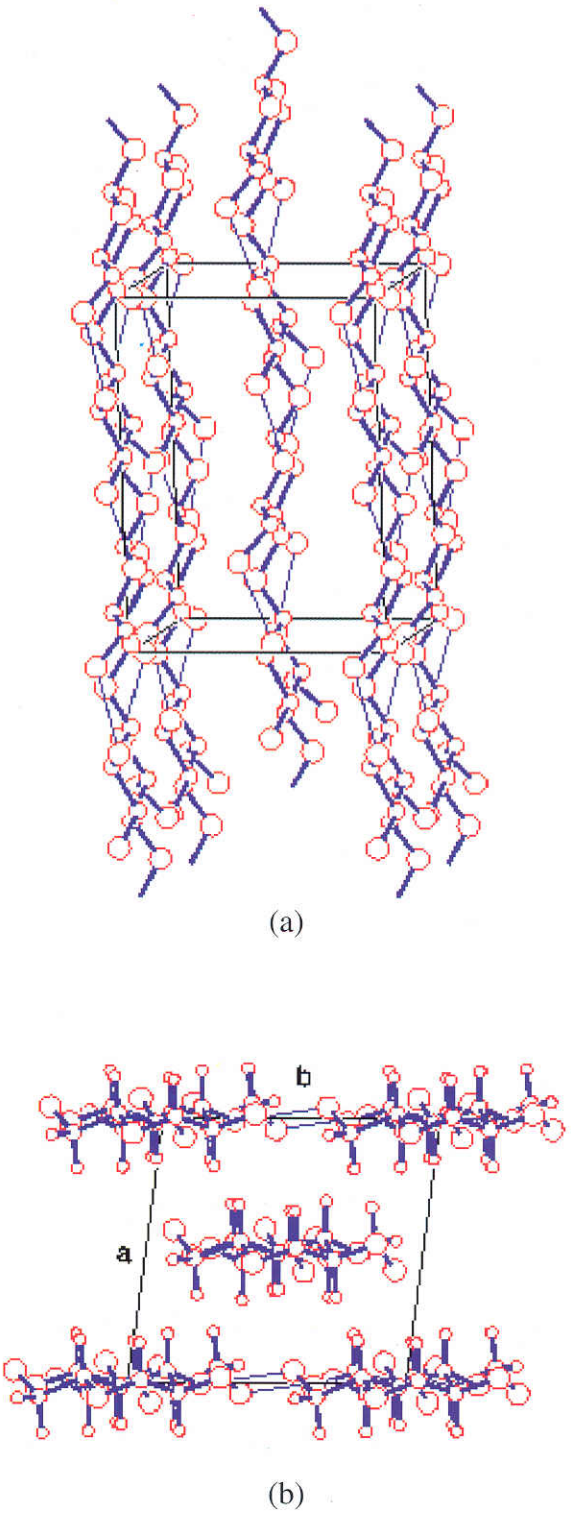
There was some discussion in the late 1980s that inconsistencies in the X-ray structures of cellulose I might be due to the paucity of the X-ray data, difficulties in accurately measuring the X-ray data, and differences between the different refinement

protocols used. However, Millane and Narasaiah showed, at least for ramie, that the WS structure was well-supported even if one considered other X-ray data sets [8], or used different refinement programs [9].

In the late 1990s, information on the up versus down packing of native cellulose began to emerge from results obtained using other techniques. Molecular mechanics and dynamics calculations showed a preference for the parallel up structure [10-12]. An intricate experiment by Koyama *et al.* in 1997 involving electron diffraction, cellobiohydralase digestion, and silver staining of an algal cellulose gave very good direct evidence for parallel up packing [13]. The lingering inconsistent X-ray results remained however.

The question of up versus down packing, and reconciliation of the X-ray results, was finally settled by Finkenstadt and Millane in 1998 by conducting a careful reanalysis of the SM and GB X-ray data [14]. Sterically flexible models of each possible packing were subjected to least-squares joint refinement against the two X-ray data sets and steric restraints. Analysis of both the X-ray agreement and the steric compression of the refined models showed that both data sets unequivocally support parallel up, and not parallel down, structures. The details can be found in Reference 14. A successful analysis of the structure is possible now (as opposed to in 1974) because refinement methods (and larger computers) available today can accommodate better joint steric/X-ray refinement of more complete and flexible molecular and crystal structure models.

The parallel up packing of native cellulose appears, therefore, to be ubiquitous throughout the plant, algal and bacterial celluloses. Views of the refined cellulose I crystal structure are shown in Fig. 1. An illuminating way of looking at the differences between up and down packing is to consider only up polarity of the chains relative to the *c*-axis, and then the down packing is obtained by changing the monoclinic angle (γ) of the unit cell from 97° to 83° [14]. To investigate the basis of the preference for parallel up over parallel down packing, cellulose chains were packed together as closely as possible without allowing steric anomalies, for different fixed values of γ [14]. The packing densities of the resulting structures were calculated as a function of γ, and the results are shown in Fig. 2. It is clear from the figure that the parallel up packing (γ = 97°) is



(a)

(b)

Figure 1: Views of the refined cellulose I crystal structure **(a)** obliquely to, and **(b)** along, the *c*-axis [14]. Thin lines show hydrogen bonds. The hydrogen atoms are excluded from (a) for clarity.

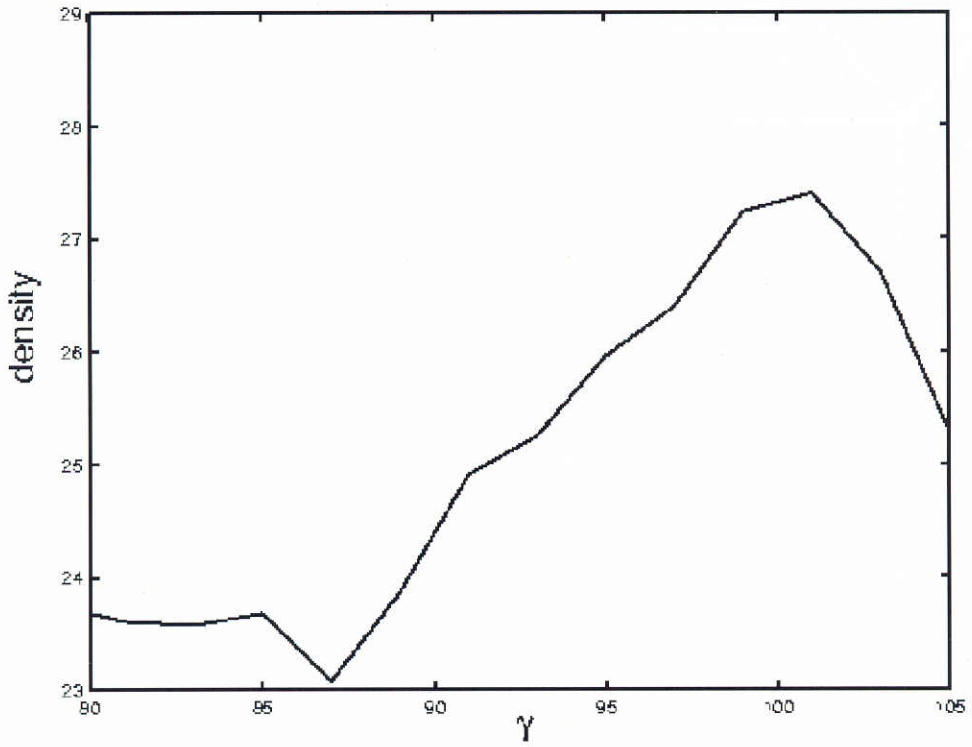


Figure 2: Packing density (arbitrary units) for sheets of cellulose molecules packed in a crystalline array as a function of the monoclinic angle γ [14].

close to the maximum packing density, whereas the parallel down packing ($\gamma = 83°$) is well away from the maximum. Of course the density so calculated is only approximate, but the trend is clear.

Is the structure of native cellulose now completely understood? Unfortunately no! In 1991, Sugiyama *et al.* [15] showed, using electron diffraction, that most celluloses are actually a mixture of two crystalline allomorphs. One of these is referred to as cellulose Ib and corresponds to that which packs in the two-chain monoclinic unit cell as described above. The other allomorph, referred to as cellulose Ia, packs in a one-chain triclinic unit cell. The two unit cells are intimately related as shown in Fig. 3, and lead to an axial shift of $c/2$ in every second sheet of cellulose molecules, between the two allomorphs. This results in a subtle difference in the packings of the two allomorphs. Algal celluloses are rich in the Ia allomorph and plant celluloses are rich in Ib. The triclinic unit cell explains the extra reflections observed in diffraction patterns from algal celluloses, that were originally attributed to a much larger monoclinic unit cell. The bimorphic structure also explains nmr data first reported in 1984 [16] that indicated the presence of two phases. The crystal structure of cellulose Ia has not been determined, but the definitive Ib structure and the relationship between the two unit cells fixes the structure in all but the fine details. There are still many questions concerning native cellulose however, including the distribution of the two allomorphs in nature, the biosynthetic mechanisms leading to the two allomorphs, and the biological function of the two subtly different packings. Finally, it is worth mentioning that Sarko and Muggli's triclinic unit cell [3], although different to that now derived from electron diffraction, did incorporate the successive $c/4$ shift of adjacent cellulose sheets more recently determined for the triclinic structure. This insightful observation 25 years ago appears to be overlooked in the modern literature.

X-ray fibre diffraction analysis of cellulose structure has a long history – spanning 60 years and continuing. From a primary structure point of view, cellulose is a simple molecule. However, it is semicrystalline and adopts a large number of

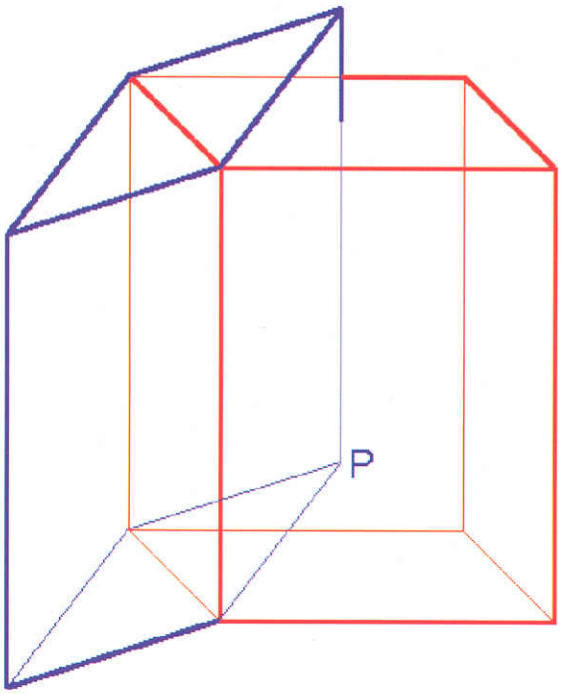


Figure 3: Relationship between the monoclinic unit cell (red) for cellulose Ib, and the triclinic unit cell (blue) for cellulose Ia. The point P has fractional coordinates (1/2,1/2,1/4) in the monoclinic system.

allomorphs. The small crystallite dimensions, inherent disorder, biphasic nature, and the presence of amorphous material mean the X-ray data from cellulose are low resolution and "murky." However, even under these conditions, X-ray fibre diffraction analysis has produced an enormous amount of essential structural information on this important material. Structural studies of cellulose demonstrate the truth of Arnott's assertion [17] that in the application of fibre diffraction "... with today's technology scrupulously applied, most gross errors are detectable."

References

[1] Meyer, K.H. and Misch, L., *Helv. Chim. Acta* (1937) **20**, 232-244.

[2] Nieduszynski, I. and Atkins, E.D.T., *Biochim. Biophys. Acta* (1970) **222**, 109-118.

[3] Sarko, A. and Muggli, R., *Macromolecules* (1974) **7**, 486-494.

[4] Gardner, K.H. and Blackwell, J., *Biopolymers* (1974) **13**, 1975-2001.

[5] Millane, R.P., in *Frontiers in Carbohydrate Research 2* (Ed. R. Chandrasekaran) 168-190 (Elsevier, London, 1992).

[6] Woodcock, C. and Sarko, A., *Macromolecules* (1980) **13**, 1183-1187.

[7] French, A.D. and Howley, P.S., in *Cellulose and Wood* (Ed C. Schuerch) 159-167 (Wiley, New York, 1989).

[8] Millane, R.P. and Narasaiah, T.V., in *Cellulose and Wood* (Ed C. Schuerch) 39-51 (Wiley, New York, 1989).

[9] Millane, R.P. and Narasaiah, T.V., *Polymer* (1989) **30**, 1763-1767.

[10] Aabloo, A., French, A.D., Mikelsaar, R.H. and Pertsin, A.J., *Cellulose* (1994) **1**, 161-168.

[11]Heiner, A.P., Sugiyama, J. and Teleman, O., *Carbohydr. Res.* (1995) **275**, 207-223.

[12] Kroon-Batenburg, L.M.J., Bouma, B. and Kroon, J., *Macromolecules* (1996) **29**, 5695-5699.

[13] Koyama, M., Helbert, W., Imai, T., Sugiyama, J. and Henrissat, B., *Proc. Natl. Acad. Sci. USA* (1997) **94**, 9091-9095.

[14] Finkenstadt, V.F. and Millane, R.P., *Macromolecules* (1998) **31**, 7776-7783.

[15] Sugiyama, J., Vuong, R. and Chanzy, H., *Macromolecules* **24** (1991) 4168-4175.

[16] Atalla, R.H. and van der Hart, D.L., *Science* (1984) **223**, 283-285.

[17] Arnott, S., in *Biopolimeri* (Associazione Italiana di Scienza e Tecnologia delle Macromolecole) Chapt. 23 (Italy, 1984).

CCP13 Software Development

M.W. Shotton[1,2], R.C. Denny[1] and V.T. Forsyth[3]

1 CLRC Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD, U.K.

2 Biophysics Section, Blackett Laboratory, Imperial College, London, U.K.

3 Institut Laue-Langevin, BP 156 F-38042, Grenoble Cedex 9, France.

There have been significant changes to two of the programs in the CCP13 suite. The CONV program has been replaced by XCONV, which is driven by an OSF/Motif-based graphical user interface (GUI). XCONV provides for the conversion of various image data files to BSL format and is aimed at being more user-friendly than CONV, especially in the case of multiple file processing. There have also been several changes to XFIX, including the